GeUmetric Deep Learning Workshop
Umeå 17-19 August 2026

Book of Abstracts

Monday 17/8

Hampus Linander (Chalmers University of Technology)

Title: Soft geometric inductive bias

Abstract: While exact equivariance can be a powerful inductive bias, real-world applications frequently exhibit approximate or broken symmetries in practice. We introduce object-centric world models that provide a soft geometric inductive bias without enforcing exact symmetries. By embedding object states as Clifford multivectors, our models encourage physically meaningful transformations while retaining the expressivity needed to handle asymmetric dynamics. We evaluate our approach on 2D rigid-body dynamics, 3D charged-particle systems, and real-world driving trajectories. Compared to both unstructured baselines and strictly equivariant models, our soft Clifford transformer achieves better long-horizon fidelity, particularly in regimes with broken symmetries. These results suggest that geometric algebra offers an effective middle ground, delivering sample-efficient dynamics without the inflexibility of hard mathematical constraints.

Jakob Galley (Umeå University)

Title: Conservation Laws from Data Symmetry in Neural Networks

Abstract: In this talk, we will explore whether intrinsic symmetries of the training data lead to conserved quantities during gradient-flow training of neural networks. We will show that under the assumption that the loss function is analytic and non-polynomial, data symmetries generically do not induce any additional integrals of motion. For mean squared error loss, on the other hand, there are situations in which data augmentation yields extra conserved quantities. I will present a framework we developed, which utilizes tensorizable networks to describe this phenomenon. Tensorizable networks are a family of architectures whose dependence on parameters and inputs can be separated using an intermediate representation. They include linear and polynomial networks, as well as Lightning Attention.

This presentation is based on joint work with Vahid Shahverdi and Axel Flinth (arxiv:2606.10913).

Isabel Dahlgren (ETH Zürich)

Title: Deep, but not difficult: identifiability and critical points for residual networks

Abstract: Residual networks, introduced in 2015, enabled the training of significantly deeper neural networks. The machine learning literature offers various explanations for their improved training properties, but most are heuristic. I will discuss a useful mathematical framework for studying residual networks — and neural networks more broadly — based on a simple idea:

replacing all activations with polynomials. Then I will present results that may help explain why deep residual networks are easier to train.

Viktor Vigren Näslund (Umeå University)

Title: PolyNODE: Variable-dimension Neural ODEs on M-polyfolds

Abstract: Neural ordinary differential equations (NODEs) are geometric deep learning models based on dynamical systems and flows generated by vector fields on manifolds. Despite numerous successful applications, particularly within the flow matching paradigm, all existing NODE models are fundamentally constrained to fixed-dimensional dynamics by the intrinsic nature of the manifold's dimension. In this paper, we extend NODEs to M-polyfolds (spaces that can simultaneously accommodate varying dimensions and a notion of differentiability) and introduce PolyNODEs, the first intrinsic variable-dimensional flow-based model in geometric deep learning. As an example application, we construct explicit M-polyfolds featuring dimensional bottlenecks and PolyNODE autoencoders based on parametrised vector fields that traverse these bottlenecks. We demonstrate experimentally that our PolyNODE models can be trained to solve reconstruction tasks in these spaces, and that latent representations of the input can be extracted.

Longde Huang (Chalmers University of Technology)

Title: Boosting Data Augmentation with Stochastic Weight Averaging

Abstract: The symmetries of a learning task have become an important factor in designing modern deep learning solutions. Data augmentation is a straightforward and effective way of incorporating symmetries into a generic neural network. Recent results show that infinitely large deep ensembles show perfect symmetry when trained on augmented data. However, since training ensembles requires repeating the training process many times, this method is costly. In this work, we study stochastic weight averaging (SWA) as an alternative ensembling technique that does not require repeated training runs. We analyze SWA by approximating the stochastic training trajectory at the end of training with an Ornstein--Uhlenbeck process. We show that in the infinite-width limit, SWA on augmented data provides an equiviariance boost that goes beyond what could be expected from the performance increase due to SWA alone. We verify our results with extensive numerical experiments on numerous models spanning computer vision and graph classification with both discrete and continuous symmetries.

Thomas Grill (KTH Royal Institute of Technology)

Title: The Boundary of Neuromanifolds of shallow MLPs

Abstract: In this talk we will introduce and investigate the boundary of the neuromanifold of shallow, fully connected neural networks with a single output and monomial activation function. This boundary plays an interesting role, because functions located on the boundary should tend to be optimizers for the training of neural networks. When the network has only two hidden neurons, we identify the boundary as the points on tangents to a well studied geometric object with explicit description, as soon as the activation degree d is greater or equals 3. When the network has k ≥ 3 hidden neurons, we describe the interior of the neuromanifold and give a concrete subset of the k-secant variety containing the boundary when d is larger than a concrete cubic bound on k. We also mention an abstract description of the boundary in this case, which uses the notion of real border rank of tensors.

Tuesday 18/8

Juan Viguera (Astra Zeneca)

Title: Generative molecular dynamics and Implicit Transfer Operators

Abstract: Molecular dynamics (MD) simulations are a key tool for studying molecular behavior, but their reliance on very small time steps makes real-world simulations computationally expensive. Generative models offer a promising way to accelerate MD by learning to sample molecular conformations and transitions directly from simulation data. These generative molecular dynamics approaches are often powered by geometric deep learning, operating on 3D point clouds while respecting permutation symmetry over atom identities and SO(3) equivariance.

In this talk, I will give a brief introduction to MD, discuss its main limitations, and outline the opportunities opened by generative molecular dynamics. I will then provide a high-level overview of recent ideas in the field, focusing on Boltzmann emulators and implicit transfer operators for small molecules. Finally, I will highlight some open challenges and current trends.

Emma Andersdotter (Umeå University)

Title: Steerable Neural ODEs on Homogeneous Spaces

Abstract: Manifold neural ODEs model data on manifolds through continuous flows generated by a vector field, instead of discrete neural network layers. This makes them well suited for learning continuous dynamics while respecting the geometry of the underlying data space. However, they only describe the evolution of points, whereas many datasets also contain features that transform under symmetries, such as vectors or tensors. We introduce steerable neural ODEs, an extension of manifold neural ODEs that evolves both positions and features in a symmetry-consistent manner. By transporting features along trajectories using parallel transport, we obtain a coupled system of ODEs: one governing the motion on the manifold and the other governing the evolution, or steering, of the associated features.

Jiri Minarcik (Carnegie Mellon/Imperial College)

Title: Untangling Surfaces via Shape and Mesh Repulsion

Abstract: Many applications in geometry processing, simulation, and fabrication require surfaces without self-intersections. We present an energy-based method for deforming arbitrary self-intersecting triangle meshes into intersection-free surfaces while preserving their shape and mesh quality. The method combines a global shape-repulsion energy with local Minkowski penalties between intersecting mesh elements. After obtaining an untangled configuration, we reverse the deformation using an IPC-constrained recovery flow while preserving the resulting isotopy class.

Vishal Ngairangbam (Karlsruhe Institute of Technology)

Title: Theory-priors for Deep Learning in Collider Phenomenology

Abstract: The Large Hadron Collider (LHC) continues to collide protons at multi-TeV energies, collecting huge amounts of high-dimensional collision data where signatures of new physics could be already present. Due to the highly enhanced pattern recognition ability of deep learning algorithms, they are quickly becoming a standard tool to excavate such signatures from moderately processed and high-dimensional representations. Considering their extreme

mathematical complexity, it is important to ascertain their behaviour under statistical outliers, as well as various phenomenological scenarios. In this talk, I'll discuss some inductive biases that naturally arise from the underlying theory. Specifically, I'll discuss how the Lorentz symmetry must be reduced for optimality under the Neyman–Pearson lemma, how one must utilise the phase space topology of decays to reduce spurious signals coming from topological obstructions, and how infra-red and collinear safe models can improve robustness of jet classification tasks.

Johanna Marie Gegenfurtner (Technical University of Denmark)

Title: The Symmetries of Three-Layer ReLU Networks

Abstract: We develop a framework for analyzing parameter symmetries in deep ReLU networks and obtain a complete characterization of the generic parameter fibers for three-layer bottleneck architectures. Our approach provides explicit semi-algebraic descriptions of these fibers and yields a polynomial time algorithm for deciding functional equivalence of two parameters. The symmetries include discrete and continuous transformations arising from layer composition, and depend on whether deeper layers hide or preserve geometric structure from preceding layers. Finally, we show that some of these symmetries induce local conservation laws along gradient flow.

(Joint work with Moritz Grillo and Guido Montúfar)

Li Ju (Uppsala University)

Title: Geometry-Aware Uncertainty Quantification for Frozen Vision–Language Models

Abstract: Vision-language models like CLIP map images and captions to deterministic points on a shared unit hypersphere S^(d-1), giving no signal of how much to trust each prediction. My recent work builds post-hoc methods that recover calibrated uncertainty from frozen encoders while respecting this geometry. AsymVLM captures aleatoric ambiguity by modeling text as directional distributions (von Mises–Fisher and power spherical) on the sphere. REPVLM estimates epistemic uncertainty through Riemannian flow matching, using embedding density as a proxy for model ignorance. GeoFlowVLM unifies both via a joint density on the product hypersphere S^(d-1) x S^(d-1). The recurring lesson: respecting the manifold (directional distributions and geodesic flows over Euclidean surrogates) is what makes uncertainty estimates calibrated and efficient. I will discuss these geometric ingredients and what they buy over geometry-blind alternatives.

Philipp Misof (Chalmers University of Technology)

Title: Finite-Width Neural Tangent Kernels from Feynman Diagrams

Abstract: The Neural tangent kernel (NTK) is a powerful tool for analyzing deep, non-linear neural networks. In the infinite-width limit the training dynamics become analytically solvable and can be expressed in terms of this kernel. However, it is well known that this limit falls short of capturing important properties such as NTK evolution or feature learning. A way to resolve this is to compute finite-width corrections to the Gaussian statistics at infinite width. In this talk, I will present a diagrammatic framework for computing these corrections to NTK statistics. These dramatically simplify the necessary algebraic manipulations and enable the computation of layer-wise recursion relations for arbitrary statistics involving preactivations, NTKs and certain higher-derivative tensors required to predict the training dynamics at leading order. Afterwards, I

will discuss how this diagrammatic approach can be used to derive new interpretable finite-width results of the NTK. To further support our findings I will present numerical evaluations of the first-order corrections and confirm close agreement to empirically sampled statistics for widths $n \gtrsim 20$.

Johan Henriksson (Umeå University)

Title: Gene function prediction - Challenges and opportunities

Abstract: Most of biology is aimed at one thing: Figuring out how things work. Here I will review how this problem is approached on the level of genes; from DNA and RNA to protein. I will make an attempt at anchoring this in geometric deep learning in how the problem of gene function prediction can, and can not, be expressed. This will likely lead the necessity of foundational models to recover useful invariants and scale separations.

Johannes Maeß (TU Berlin)

Title: Implicit Machine Learning Force Fields: Fast Dynamics for Molecular Geometries

Abstract: Learned models now push the accuracy–efficiency frontier in simulating physical systems — yet they are still applied statelessly: each step is computed from scratch, despite the simulated system evolving continuously. Machine learning force fields (MLFFs) driving molecular dynamics rebuild the latent representation of a molecular geometry at every femtosecond step, even though its atoms have moved by less than a hundredth of an atomic bond length.

We remove this redundancy by defining the representation implicitly. Instead of stacking distinct message-passing layers to make a prediction, we iterate a single layer to self-consistency and take its fixed-point solution as the representation of the molecular geometry. As the geometry of the molecule evolves smoothly along a trajectory, so does this representation — which makes it predictable. Extrapolating it forward from previous timesteps gives the solver a starting point close to the answer, reducing solving cost to just one to two iterations, with more spent automatically where the geometry changes fastest.

The fixed-point formulation further allows implicit differentiation to yield energy-conserving forces, which slashes memory consumption down to that of a single message-passing layer. Across invariant, Cartesian-tensor and SO(3)-equivariant architectures this gives a two- to five-fold reduction in compute and memory. I-MLFFs retain the atomistic resolution, fine integration timestep, and quantum-chemical accuracy of deep explicit architectures, at the footprint of a shallow single-layer model — allowing for the simulation of larger systems on the same budget and GPU. I will close with how I-MLFFs can improve modelling of long-range effects that happen on larger molecules, and how they generalize to other physical simulations, in the example of implicit Hamiltonian neural networks.

Wednesday 19/8

Andreas Ipp (TU Wien)

Title: Lattice Gauge-Equivariant Neural Networks: From Fixed-Point Actions to Diffusion Models

Abstract: Quantum field theories are governed by local gauge symmetries. On the lattice, they correspond to local gauge transformations that act independently at every point of a discretized spacetime. Representing these local non-Abelian symmetries in deep neural architectures presents unique challenges compared to traditional global geometric equivariance. In this talk, I introduce Lattice Gauge Equivariant Convolutional Neural Networks (L-CNNs), a framework explicitly designed to preserve exact local gauge equivariance on discretized spacetimes using non-Abelian parallel transporters and Wilson loops.

Building on this theoretical foundation, I will showcase two recent applications of the L-CNN architecture. First, I present machine-learned renormalization-group-improved gauge actions, where gauge equivariance enables the parameterization of classically perfect fixed-point actions that suppress discretization artifacts even on coarse grids. Second, I discuss gauge-equivariant diffusion models, where incorporating local symmetries into score-based generative frameworks leads to highly generalizable sampling dynamics for non-Abelian lattice gauge theories. Together, these results demonstrate how L-CNNs bridge fundamental geometric principles of gauge theories with practical deep learning models for physical simulations.

Miaowen Dong (Chalmers University of Technology)

Title: Equivariance and Augmentation for Bayesian Neural Networks

Abstract: Symmetries are important for many deep learning tasks, ranging from applications in the sciences to medical imaging. However, there is an ongoing debate about whether to impose symmetry constraints on the neural network architecture (yielding equivariant neural networks) or learn them from augmented training data. Although equivariant networks are well-studied theoretically, much less is known about data augmentation, since analyzing augmentation requires control over the training dynamics. Inspired by recent results that show that augmented infinite deep ensembles are exactly equivariant, we study data augmentation for Bayesian neural networks (BNNs) trained with variational inference. We focus on variational distributions in the exponential family and derive conditions under which exact equivariance is reached. We furthermore obtain bounds on the equivariance error and introduce three novel symmetrization techniques which boost the effect of data augmentation in this setting. We conduct extensive numerical experiments which show that one of our symmetrization methods (orbit expansion) outperforms the baseline in both equivariance and overall performance.

Aron Persson (Umeå University)

Title: Stochastic currents as diffusion laws

Abstract: We apply stochastic currents to the setting of diffusion laws in the context of machine learning using scores. An outline to how this connects to empirical samples and give some consistency results. Lastly we give an outlook on generalizing stochastic currents to M-polyfolds.

Oguzhan Recep Akkol (Aix-Marseille University)

Title: Detecting ADE Singularities via Realization Fibers: Toward Explainable Polynomial Neural Networks

Abstract: Understanding what a trained network's internal structure encodes is a central question in explainability, and singularity type provides a well-studied and natural target through which to investigate this question. First, we study the exact representation of ADE-family hypersurface singularities by polynomial neural networks with monomial activation functions. Second, we consider a specially constructed architecture that enables us to detect algebraic invariants of the singularities of the target polynomial. We show that there is an exact relationship between the Milnor number of the target polynomial and the fiber kernel dimension, which we verify computationally across a range of cases. Finally, we connect these findings to Singular Learning Theory by relating this algebraic detection mechanism to the geometry of flat directions in the loss landscape. This provides a step toward explainability for polynomial neural networks.

This is joint work with Meral Tosun, Alptug Batuhan Softa, and Naki Eren Sengezer.